

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/280042132>

# Basis Pursuit Receiver Function

Article in *Bulletin of the Seismological Society of America* · November 2014

Impact Factor: 2.32 · DOI: 10.1785/0120140004

---

READS

83

4 authors:



[Mrinal K. Sen](#)

University of Texas at Austin

213 PUBLICATIONS 3,310 CITATIONS

[SEE PROFILE](#)



[Reetam Biswas](#)

University of Texas at Austin

5 PUBLICATIONS 2 CITATIONS

[SEE PROFILE](#)



[Prantik Mandal](#)

National Geophysical Research Institute

87 PUBLICATIONS 1,086 CITATIONS

[SEE PROFILE](#)



[P. Kumar](#)

National Geophysical Research Institute

53 PUBLICATIONS 1,297 CITATIONS

[SEE PROFILE](#)

## Basis Pursuit Receiver Function

by Mrinal K. Sen,<sup>\*</sup> Reetam Biswas,<sup>\*</sup> Prantik Mandal, and Prakash Kumar

**Abstract** Receiver functions (RFs) are derived by deconvolution of the horizontal (radial or transverse) component of ground motion from the vertical component, which segregates the *PS* phases. Many methods have been proposed to employ deconvolution in frequency as well as in time domain. These methods vary in their approaches to impose regularization that addresses the stability problem. Here, we present application of a new time-domain deconvolution technique called basis pursuit deconvolution (BPD) that has recently been applied to seismic exploration data. Unlike conventional deconvolution methods, the BPD uses an L1 norm constraint on model reflectivity to impose sparsity. In addition, it uses an overcomplete wedge dictionary based on a dipole reflectivity series to define model constraints, which can achieve higher resolution than that obtained by the traditional methods. We demonstrate successful application of BPD based RF estimation from synthetic data for a crustal model with a near-surface thin layer of thickness 5, 7, 10, and 15 km. The BPD can resolve these thin layers better with much improved signal-to-noise ratio than the conventional methods. Finally, we demonstrate application of the BPD receiver function (BPRF) method to a field dataset from Kutch, India, where near-surface sedimentary layers are known to be present. The BPRFs are able to resolve reflections from these layers very well.

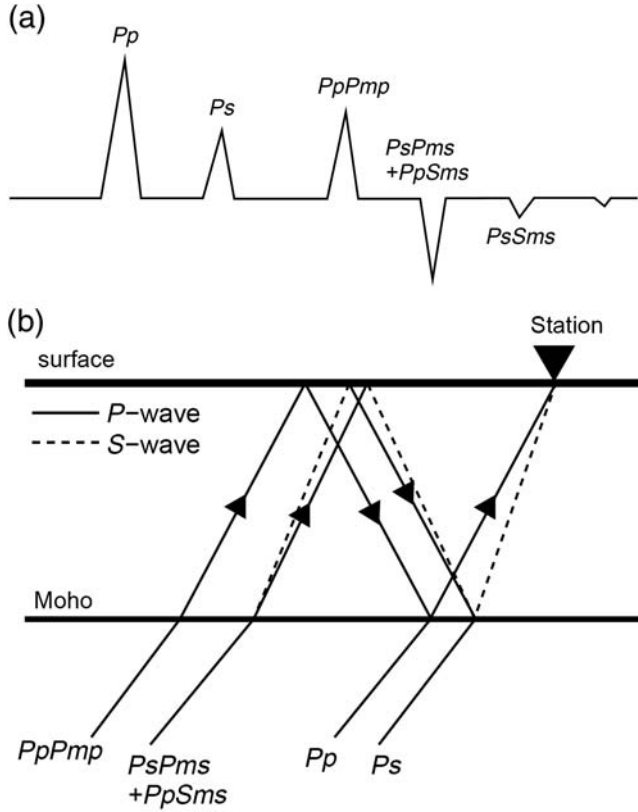
### Introduction

It has now become a common practice to image crust and upper-mantle discontinuities using receiver functions (RFs). An RF is essentially a teleseismic record of *PS* conversion without the complexity of the source time function. [Vinnik \(1977\)](#) demonstrated one of the earliest applications of RF to image the 410 and 660 km discontinuities using *PS* converted waves. A seismic trace recorded at a station contains information about the source structure, mantle propagation effects, and the crustal structure beneath the station. Figure 1 shows a ray interpretation of the Earth's layer response to an incoming plane wave ([Ammon, 1991](#)). At near normal incidence (for example, at teleseismic distances), the *S* waves are mostly recorded in the horizontal-component geophones and are negligible in the vertical receivers. The RFs, which are generated by deconvolution of the radial or transverse component of the seismogram from the vertical component, contains *PS* converted phases and *S*-phase reverberations. The process of deconvolution isolates the near-receiver propagation effects from the far-field dislocation time function, near-source structure, and lower-mantle propagation effects. Thus, the RFs can be used to image crust and upper-mantle structures beneath the recording station (e.g., [Phinney, 1964](#); [Burdick and Langston, 1977](#)).

Several variants of RF deconvolution have been proposed and implemented in frequency and time domains. Some algorithms are based on iterative least-squares inversion techniques, and others calculate the direct inverse of the problem. [Phinney \(1964\)](#) used frequency domain deconvolution for *P*-waveform modeling of spectral responses from the crust at several Worldwide Standard Seismograph Network (WWSSN) stations. There exist several problems associated with frequency domain deconvolution mainly due to band limitation of the recorded data. To address this, [Burdick and Langston \(1977\)](#) introduced the time-domain deconvolution for RF estimation. The time-domain inversion helps to better preserve the phase information of the signal. [Ammon \(1991\)](#) enhanced it further by proposing a method to estimate the absolute amplitude of RFs and also the variation of amplitude of converted phases with source–receiver distance. He also showed the importance of estimation of true amplitude as it gets affected by scattering, which may be indicative of anisotropy of the structure below. [Gurrola et al. \(1995\)](#) used a simultaneous time-domain deconvolution method, which uses various norms, and were able to improve on the results by suppressing side lobes and increasing the resolution of the peak phases.

In this paper, we use a compressive sensing based technique, namely, the basis pursuit deconvolution (BPD) for RF deconvolution in time domain. The BPD is based on the idea

<sup>\*</sup>Now at the Institute for Geophysics, The University of Texas at Austin, 10100 Burnet Road, Building 196, Austin, Texas 78730.



**Figure 1.** (a) A plot showing arrivals of various direct and converted phases of teleseismic  $P$  phases on the radial receiver function (RF). (b) Ray paths of  $P$ -to- $S$  conversions associated with the Moho and other crustal multiples recorded by a three-component seismograph considering crust overlying an upper mantle for an incoming teleseismic  $P$  wave.

of using an L1 norm minimization technique (Chen *et al.*, 1998), which has recently been applied to seismic exploration problems (Puryear and Castagna, 2008; Zhang and Castagna, 2011; Zhang *et al.*, 2013). For optimization of our objective function, we use the Gradient Projection Sparse Reconstruction (GPSR) technique with theoretically computed regularization weights that are iteration adaptive (Figueiredo *et al.*, 2007; Biswas and Sen, 2014). We show examples of BPD based RF estimation for several synthetic datasets and a real dataset from one three-component seismograph station of the local seismic network of the National Geophysical Research Institute, Hyderabad, in the Kutch rift zone, Gujarat, India.

### Theory

The fundamental issue in the generation of an RF is the deconvolution; the problem associated with deconvolution is fairly well known. Several approaches have been proposed by the signal processing and seismology communities to address these issues. In an RF, the recorded vertical component trace is assumed to be the source wavelet containing mantle and near-source propagation effects. This is convolved with the impulse response (propagation effects below the receiver) to generate the horizontal (or radial) component. The process

can be represented by the following simple convolution model:

$$H(t) = r(t) * V(t), \quad (1)$$

in which  $r(t)$  is the earth's response. Because the vertical component  $V(t)$  and horizontal component  $H(t)$  both have the source and mantle propagation effects, on deconvolution they get eliminated and we generate the  $PS$  converted phases (Fig. 1). A similar procedure is employed to generate an  $S$ -RF in which the radial component is deconvolved from the vertical component (Ligorria and Ammon, 1999). All our discussions will apply to both  $P$ - and  $S$ -RFs unless otherwise stated although we will discuss  $P$ -RFs only.

### Existing Deconvolution Methods

Mostly RFs are calculated in the frequency domain by spectral division (e.g., Vinnik, 1977; Owens *et al.*, 1984) together with some regularization so as to avoid instability. In frequency domain it takes the following form:

$$R(\omega) = \frac{H(\omega)V^*(\omega)}{V(\omega)V^*(\omega) + \delta}, \quad (2)$$

in which  $R$  is the Fourier transform of the RF,  $H$  and  $V$  are the Fourier transforms of horizontal and vertical component, respectively,  $*$  represents the complex conjugate, and  $\delta$  represents the added white noise. The white noise is added to damp the solution and thus it stabilizes it against small values by prewhitening the denominator. The above formulation can be modified to carry out deconvolution for multiple traces simultaneously instead of stacking results for individual traces. The above formulation for simultaneous damping factor deconvolution can be written as

$$R(\omega) = \frac{\sum H(\omega)V^*(\omega)}{\sum V(\omega)V^*(\omega) + \delta}, \quad (3)$$

in which the summation is performed over the number of traces. The prewhitening or damping can also be carried out by water-level deconvolution (Owens *et al.*, 1984), which is similar to damping factor deconvolution but instead of adding a constant factor  $\delta$  to the denominator at all frequencies, it employs damping only at those frequencies at which the values fall below a certain fraction of the peak amplitude.

In the two methods described above, the value of  $\delta$  is selected by trial and error. In the damped least squares deconvolution method employed by Bostock (1998), the value of  $\delta$  is selected by regression analysis through minimization of the general cross validation (GCV) function, which is defined by

$$GCV(\delta) = \frac{\sum_M \sum_N (H - VR)^2}{(MN - \sum_N R)}, \quad (4)$$

in which  $N$  is the number of frequencies, and  $M$  is the number of traces.

All the methods described above use a constant value for  $\delta$ . [Chen et al. \(2010\)](#) proposed an array-conditioned deconvolution method, which is data adaptive and independently determines  $\delta$  for each frequency. This technique gives better results in multichannel processing than simultaneous deconvolution, which is biased toward single channel, multievent data ([Chen et al., 2010](#)). The filter derived in this method minimizes the noise energy and energy related to instability and thus maximizes the signal energy ([Haldorsen et al., 1994, 1995](#)).

Apart from the frequency domain spectral division there are several time-domain deconvolution methods. The time-domain methods mainly solve an optimization problem derived from equation (1), which can be represented by the following linear system of equations

$$\mathbf{d}^{\text{syn}} = \mathbf{G}\mathbf{m}, \quad (5)$$

in which  $\mathbf{d}^{\text{syn}}$  represents a synthetic horizontal component, the matrix  $\mathbf{G}$  represents the operator matrix containing shifted vertical seismogram representing a convolution kernel, and  $\mathbf{m}$  represents the vector of the model RF.

Here, we try to minimize the misfit between the synthetic and the observed horizontal seismograms. Similar to the procedure employed in the frequency domain, the system is solved with damping or constraints using the least-squares approach. [Gurrola et al. \(1995\)](#) applied a time-domain technique in simultaneous deconvolution, in which instead of solving for an individual RF for each trace and stacking all the ones for a small range of distance and back azimuth, it solves for all traces simultaneously.

Typically, the optimization problem solves for unknown model parameters by minimizing a functional that contains a data misfit term and a model norm, in which each uses an L2 norm. The objective function is given by

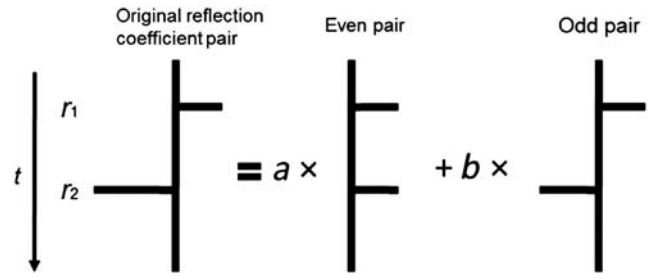
$$\|\mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{syn}}\|_2^2 + \lambda \|\mathbf{m}\|_2^2, \quad (6)$$

in which  $\mathbf{m}$  represents the RF vector and  $\lambda$  is a regularization weight that determines relative importance of data and model norms.

Some other variants of RF estimation were proposed by [Kikuchi and Kanamori \(1982\)](#) and [Ligorria and Ammon \(1999\)](#), who solved the RF deconvolution in time domain by a forward iterative deconvolution technique. This method measures the similarity of two waveforms by a cross-correlation function. Here, the vertical and horizontal components of the seismograms are cross correlated, and a peak in RF series is placed at the point of maximum cross correlation. This RF series is then convolved with the vertical component to generate the predicted trace, which is then subtracted from the observed trace, and this process is iterated until a certain misfit is reached.

### Basis Pursuit Deconvolution

One common approach to reflectivity estimation is to replace the L2 model norm in equation (6) with an L1 norm.



**Figure 2.** Any reflection pair can be represented as the sum of linear combination of even and odd dipole pairs. Thus, the coefficients  $a$  and  $b$  are now the new model parameters ([Zhang et al., 2013](#)).

The mixed L1–L2 norm poses some difficulties in solving the system, but it introduces sparsity in the reflectivity series (RF), reducing unwanted side lobes. The BPD uses a mixed norm objective function problem having L1 norm of the model and L2 norm of the data error, which introduces sparsity in layers ([Zhang and Castagna, 2011; Zhang et al., 2013](#)). We describe the method in brief; however complete details can be found in [Zhang and Castagna \(2011\)](#). The BPD attempts to minimize the L2 norm of data jointly with the L1 norm of the model parameters using a regularization parameter ( $\lambda$ ) that determines the relative weight between the two norms. The objective function is given by

$$\begin{aligned} E(\mathbf{m}) &= \|\mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{syn}}\|_2^2 + \lambda \|\mathbf{m}\|_1 \\ &= \|\mathbf{d}^{\text{obs}} - \mathbf{G}\mathbf{m}\|_2^2 + \lambda \|\mathbf{m}\|_1. \end{aligned} \quad (7)$$

In a Bayesian framework, the result of an inverse problem is described by the posterior probability density function (PPD) in model space, which is the product of a likelihood function and an *a priori* probability density function (PDF) in model space (e.g., [Tarantola, 1984; Sen and Stoffa, 2013](#)). In the present context with reference to equation (7), the PPD is therefore proportional to  $\exp(-E(\mathbf{m}))$ . Thus, the first part (or the data norm) is the exponent of the likelihood function, which is a Gaussian, and the second part, that is, the model norm corresponding to the prior PDF, is a Laplace distribution.

In BPD, we decompose the reflectivity series in terms of a linear combination of even ( $r_e$ ) and odd ( $r_o$ ) dipoles (Fig. 2). Using these dipole pairs, we construct a dipole dictionary, termed the wedge dictionary by [Zhang and Castagna \(2011\)](#), so as to accommodate layers of varying thickness arranged with increasing time separation  $\Delta t$ . The even and odd dipoles are given by

$$\begin{aligned} r_e(t, p, q, \Delta t) &= \delta(t - p\Delta t) + \delta(t - p\Delta t + q\Delta t); \\ r_o(t, p, q, \Delta t) &= \delta(t - p\Delta t) - \delta(t - p\Delta t + q\Delta t), \end{aligned} \quad (8)$$

in which  $p$  represents the shift of reflectivity pair in time and varies from 1 to  $P$ , total samples in the trace; and  $q$  represents the bed time thickness, which ranges from 1 to maximum bed time thickness  $Q$ . The dipole dictionary when multiplied with the corresponding coefficients  $\mathbf{a}$  and  $\mathbf{b}$  produces the

reflectivity series. Mathematically, the reflectivity series in terms of the dictionary can be represented as

$$r(t, p, q, \Delta t) = \sum_{p=1}^P \sum_{q=1}^Q [a_{pq} r_e(t, p, q, \Delta t) + b_{pq} r_o(t, p, q, \Delta t)]. \quad (9)$$

The dipole dictionary is convolved with the wavelet  $\mathbf{W}$  (vertical component) resulting in kernel matrix  $\mathbf{K}$  replacing  $\mathbf{G}$  in equation (5). This gives rise to the following new equation:

$$\mathbf{d}^{\text{syn}} = \mathbf{K}\mathbf{h}, \quad (10)$$

in which the model vector  $\mathbf{h} = [\mathbf{a} \ \mathbf{b}]^T$  contains a series of coefficients  $\mathbf{a}$  and  $\mathbf{b}$ . Also, the seismogram (radial component) obtained by our forward model equation (1) can be expressed in terms of the kernel matrix as

$$H(t) = \sum_{p=1}^P \sum_{q=1}^Q [a_{pq} W * r_e(t, p, q, \Delta t) + b_{pq} W * r_o(t, p, q, \Delta t)] = \mathbf{K}\mathbf{h}. \quad (11)$$

Thus our objective function changes to the following form:

$$E(\mathbf{h}) = \|\mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{syn}}\|_2^2 + \lambda \|\mathbf{h}\|_1 = \|\mathbf{d}^{\text{obs}} - \mathbf{K}\mathbf{h}\|_2^2 + \lambda \|\mathbf{h}\|_1, \quad (12)$$

which is solved for the coefficients of even and odd dipoles. We have used the GPSR with theoretical lambda calculation to solve for the optimal value of  $E(\mathbf{h})$ . Because of the representation of RF in terms of sparse basis, its resolution is expected to increase (Zhang and Castagna, 2011).

#### Gradient Projection Sparse Reconstruction with Theoretical Lambda

GPSR (Figueiredo *et al.*, 2007) is a recently developed L1 norm minimization algorithm, which solves our unconstrained optimization problem described by equation (7) or (12) by converting it to a quadratic formulation. Algorithm details can be found in Figueiredo *et al.* (2007). We will describe this method briefly, for completeness. The GPSR first defines two vectors  $\mathbf{u}$  and  $\mathbf{v}$  such that

$$\mathbf{h} = \mathbf{u} - \mathbf{v}, \quad \mathbf{u} \geq 0, \quad \mathbf{v} \geq 0, \quad (13)$$

in which  $u_i = (h_i)_+$  and  $v_i = (h_i)_-$  for all  $i = 1, 2, \dots, n$  and  $(\cdot)_+$  denotes positive part operator. Thus,  $\|\mathbf{h}\|_1 = \mathbf{1}_n^T \mathbf{u} + \mathbf{1}_n^T \mathbf{v}$ , in which  $\mathbf{1}_n = [1, 1, \dots, 1]^T$  is the vector consisting of  $n$  ones. Thus, equation (7) can be rewritten in the form of bound-constrained quadratic program (BCQP) as follows:

$$\min_{\mathbf{u}, \mathbf{v}} \frac{1}{2} \|\mathbf{d}^{\text{obs}} - \mathbf{K}(\mathbf{u} - \mathbf{v})\|^2 + \lambda \mathbf{1}_n^T \mathbf{u} + \lambda \mathbf{1}_n^T \mathbf{v}. \quad (14)$$

This can be further written in the more standard BCQP form and then solved following a gradient projection algorithm described in Figueiredo *et al.* (2007).

Biswas and Sen (2014) modified the GPSR algorithm to calculate the regularization parameter  $\lambda$  at each iteration by an iteration adaptive analytical formula using the Bayesian approach for the mixed norm (Alliney and Ruzinsky, 1994). It calculates the optimal value of  $\lambda$  at each iteration by minimizing a function of an auxiliary variable  $\alpha$  given by

$$S(\alpha) \left( \mathbf{I} + \frac{1}{\alpha} \mathbf{K}^T \mathbf{K} \right)^{1/N}, \quad (15)$$

in which  $\mathbf{I}$  is an identity matrix,  $N$  is the number of data points, and the function  $S(\alpha)$  is given by

$$S(\alpha) = \|\mathbf{z}\|_2^2 - \mathbf{z}^T \mathbf{K} [\alpha \mathbf{I} + \mathbf{K}^T \mathbf{K}]^{-1} \mathbf{K}^T \mathbf{z}, \quad (16)$$

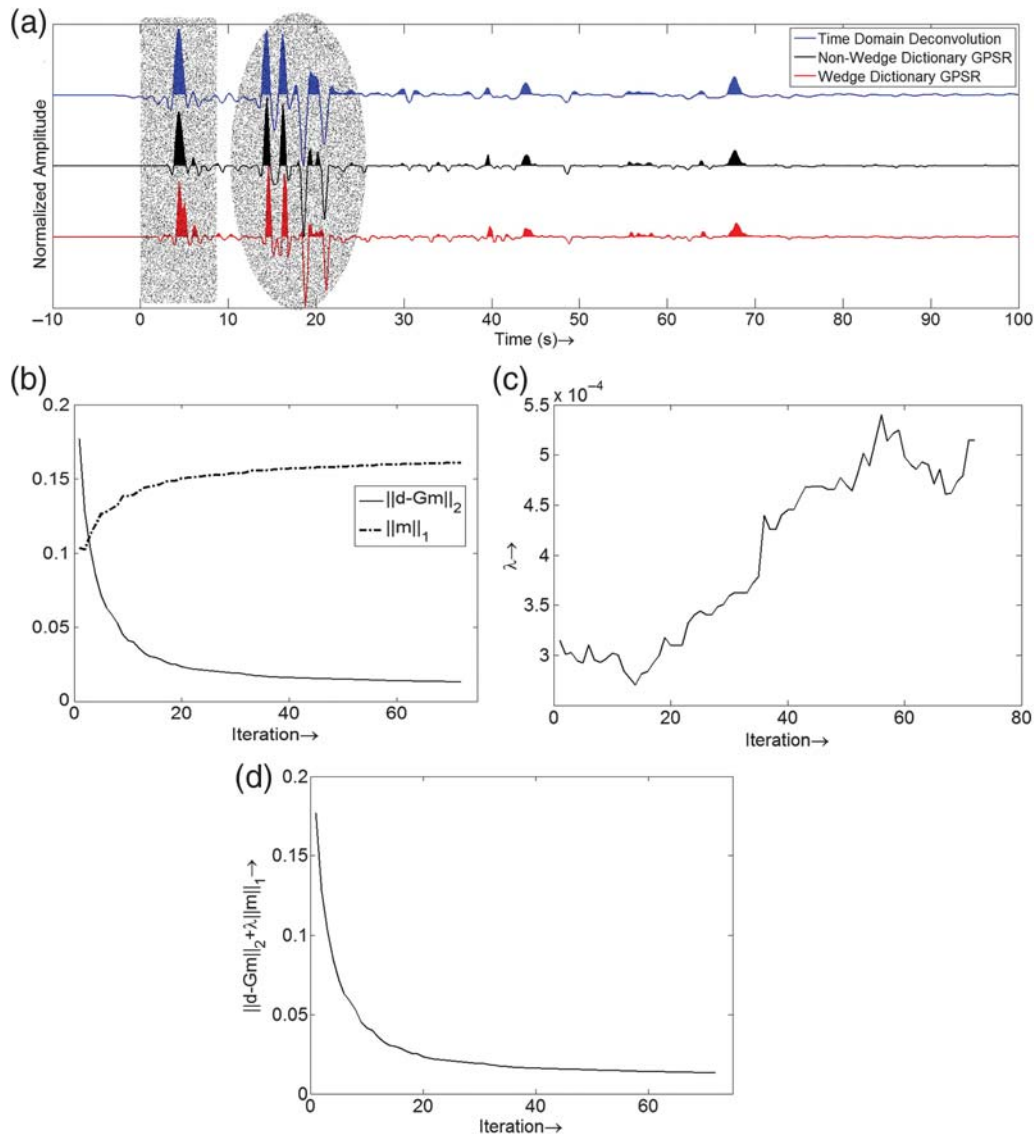
with  $\mathbf{z} = -\mathbf{d}^{\text{obs}}$  and  $\alpha = \lambda^2 \sigma^2$ , in which  $\sigma$  is the variance of the distribution.

## Examples

### Synthetic Examples

The objective of the numerical experiments reported in this section is to demonstrate the efficacy of our basis pursuit receiver function (BPRF) to resolve thin layers on a suite of synthetic data for four different crustal models each containing a thin bed (Figs. 3a and 4a–c). For this purpose, we generated synthetic vertical and radial seismograms for a simple crustal model containing a thin surface layer above Moho with increasing thickness of 5, 7, 10, and 15 km and computed synthetic seismograms using the reflectivity method (Fuchs and Müller, 1971). For each case, we generated the full waveform synthetics at approximately ~0.5 Hz dominant frequency. The plane containing the vertical and radial components is rotated such that the vertical component is along the direction of the maximum polarization direction of the  $P$  wave. This rotation results in a clean decomposition of  $P$ – $SV$ – $SH$  components. The rotated components are then subjected to time-domain deconvolution and our modified GPSR algorithm with a wedge dictionary (equation 9) and nonwedge dictionary (equation 7). Figures 3a and 4a–c show the RFs for 5, 7, 10, and 15 km thick layers, respectively. Figure 3b–e shows some characteristic curves; Figure 3b,c, and e shows the variation of data norm, model norm, and total norm, respectively, as a function of iteration. The model norm increases with iteration initially (Fig. 3c) and then becomes nearly flat beyond iteration 30. This is because our starting model contains a series of zero reflectivity (the sparsest set); the algorithm finds the minimum number of spikes or layers needed to fit the data adequately. Figure 3d shows variation of theoretically calculated  $\lambda$  as a





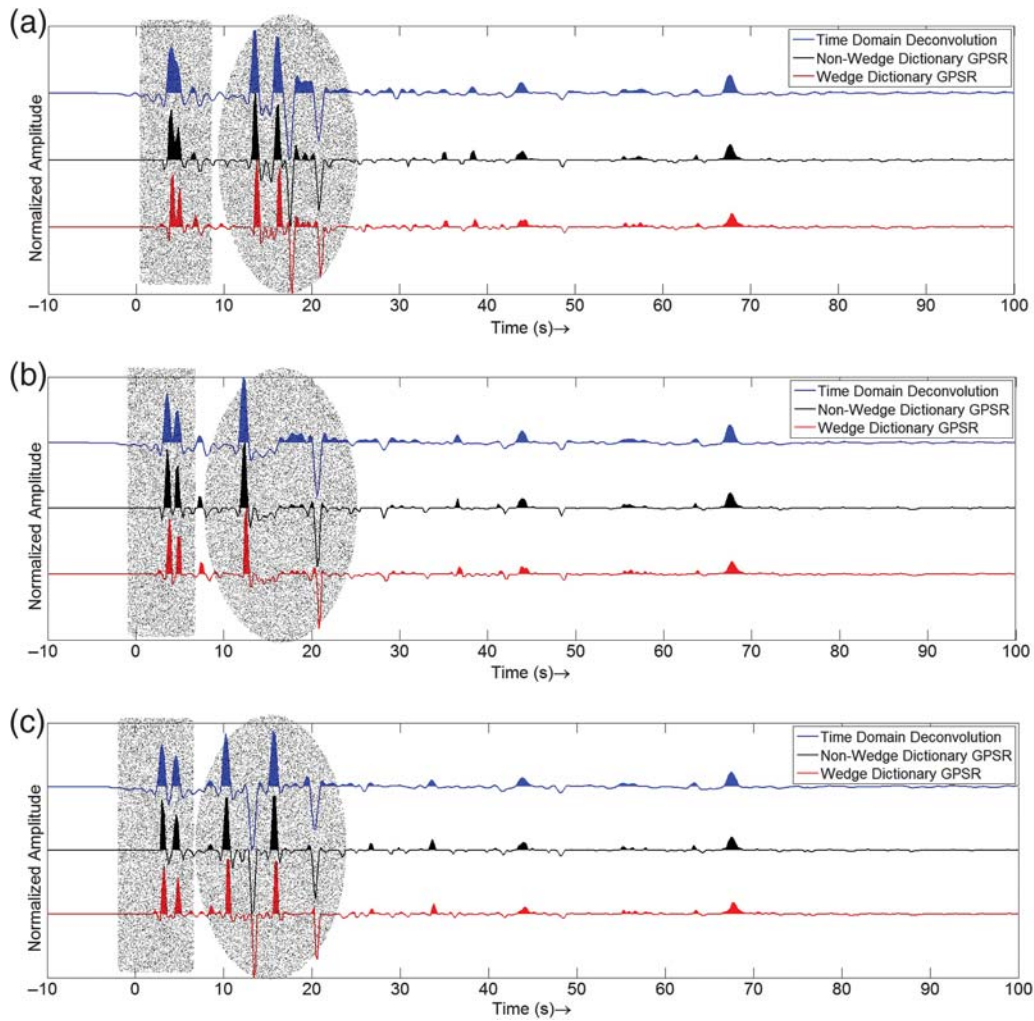
**Figure 3.** Plot of RF and other characteristic plots for a 5 km thick layer synthetic model. (a) Plot of RF for a case of a 5 km thick layer using time-domain deconvolution and Gradient Projection Sparse Reconstruction (GPSR) with nonwedge and wedge dictionary. The  $Pp$  phase has been muted to highlight the  $PS$  converted phases. Initial  $PS$  phases shown by a rectangular box are resolved only by GPSR using wedge. (b) Plot of error and model norm along with iteration for GPSR algorithm. (c) Plot of  $\lambda$  along with iteration for GPSR algorithm. (d). Plot of objective function along with iteration for GPSR algorithm.

function of iteration, which demonstrates that the variable  $\lambda$  becomes stable after about 30 iterations beyond which the error does not change significantly. It is clear that our modified GPSR algorithm with wedge dictionary was able to resolve even the 5 km thin layer case, which was not possible with the traditional method (see the peaks in the highlighted rectangles). The traditional method could not even resolve the 7 km thick layer, which was resolved by the nonwedge dictionary case (Fig. 4a). The traditional method started resolving from the 10 km (Fig. 4b) thick layer onwards. Also the later phases (highlighted by ellipses in the figures), are much better resolved by the GPSR using the wedge dictionary.

We further demonstrate that, like RFs estimated by other deconvolution techniques, BPRF can also be inverted to derive

realistic models. In Figures 5–7, we show the inversion results along with comparison with the time-domain RF for different synthetic cases. The inversion results are shown in each subplot (c). It is clearly seen that the traditional and BPRF produce almost similar models. Here, the model in Figure 5 is the same as that used in Figure 4 for the thin bed at 7 km above the Moho. The example in Figure 6 is for the bed 15 km above the Moho, whereas in Figure 7 we took the 5 km thick layer at the surface so as to mimic the real field example, which is discussed later.

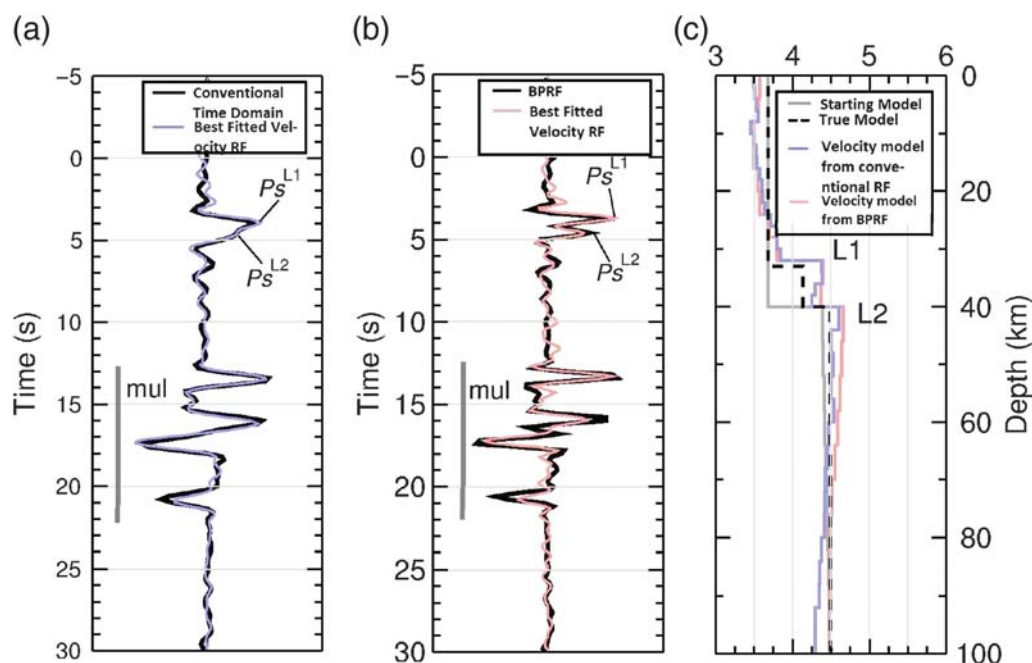
Now, let us discuss Figure 5, in which we have a 7 km thin bed; similar discussion and interpretation will apply to Figures 6 and 7. Recall that the synthetic full waveforms were generated for a model with a crust that is 40 km thick and contains a thin layer of 7 km thickness just above the



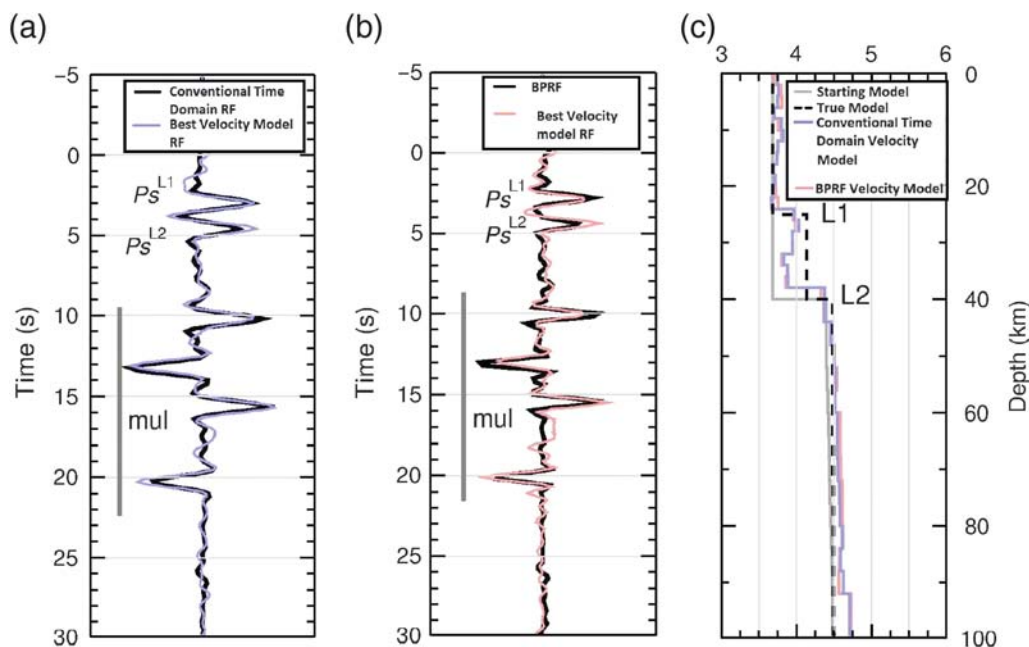
**Figure 4.** Plot of RF using time-domain deconvolution and GPSR with wedge and nonwedge dictionary for (a) 7 km, (b) 10 km, and (c) 15 km thick thin layer at the surface synthetic model. The  $Pp$  phase has been muted to highlight the  $PS$  converted phases. The resolution of the initial phases has been marked by the rectangular block and the later phases or multiples by the ellipse.

Moho. The vertical and radial waveforms are processed to estimate the traditional RF and BPRF as described in an earlier section. Note that the traditional RF (Fig. 5a) hardly sees the primary conversions (marked as  $Ps^{L1}$ , and  $Ps^{L2}$ ) from the two discontinuities, however, it produces an inverted model (blue) that is close to the true one (black dashed line). This is primarily due to the timings and amplitudes of its multiples at  $\sim 16$  and  $21$  s. The noise-free synthetics can clearly show these multiple arrivals, but real data may not record these faithfully due to noise and may also be interpreted as two different discontinuities unless and until we have good moveout curves for these. On the other hand, the BPRF (Fig. 5b) can clearly resolve the two closely separated discontinuities giving an unambiguous impression about the derived model. Figures 6 and 7 represent cases where we have well-separated discontinuities. In one case, the thin layer above the Moho is 15 km thick (Fig. 6), and in the other case a 5 km thin layer is at the top of the model (Fig. 7). Here, both the RFs clearly distinguish and resolve the  $L1$  and  $L2$  interfaces.

In the examples so far, our models consisted of horizontal layers. Next, we considered a dipping layer model and generated RFs using basis pursuit and compared those with the traditional RF (Fig. 8). First, we generated synthetic waveforms (Frederiksen and Bostock, 2000) for a model shown in Figure 8c, where both Moho and a thin layer of 10 km thickness above the Moho dip to the right. The synthetic data were generated in all back-azimuthal directions from  $0^\circ$  to  $360^\circ$  with a step of  $10^\circ$ . The dips for Moho and the layer are  $5^\circ$  and  $10^\circ$ , respectively. The generated radial and transverse data were then used to estimate the traditional and basis pursuit RFs. Figure 8a,b shows the  $SV$  and  $SH$  components, respectively, in which the traditional RFs and BPRF are superimposed to compare their amplitudes and timings. It is clearly seen that traditional RF and BPRF match well for both the  $SV$  and  $SH$  components. However, within the time window of 2–6 s (Fig. 8a,b) the BPRFs are able to resolve the thin bed that is also reflected in the  $SH$  component (b).



**Figure 5.** Synthetic RF and the inversion result for a 7 km thick layer above Moho. (a) Traditional RF for a case of 7 km thick layer using time-domain deconvolution (black) and the RF calculated from the best-fitted velocity model (blue) in panel (c). (b) RF estimated using GPSR (black), and the best-fitted velocity model is shown in panel (c) as red. (c) Velocity models calculated from the time-domain RF (blue) and BPD receiver function (BPRF) (red) along with true model (dotted black) and starting model (gray).



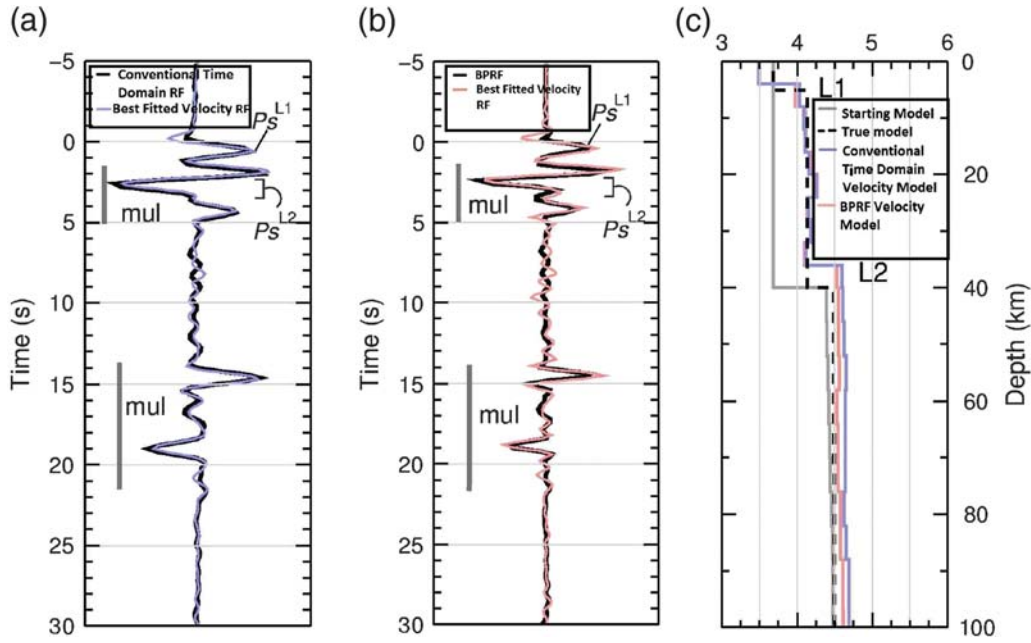
**Figure 6.** Synthetic RF and the inversion result for a 15 km thick layer above Moho. (a) Traditional RF for a case of 15 km thick layer using time-domain deconvolution (black) and the RF calculated from the best-fitted velocity model (blue) in panel (c). (b) RF estimated using GPSR (black), and the best-fitted velocity model is shown in panel (c) as red. (c) Velocity models calculated from the time-domain RF (blue) and BPRF (red) along with true model (dotted black) and starting model (gray).

#### Real Data

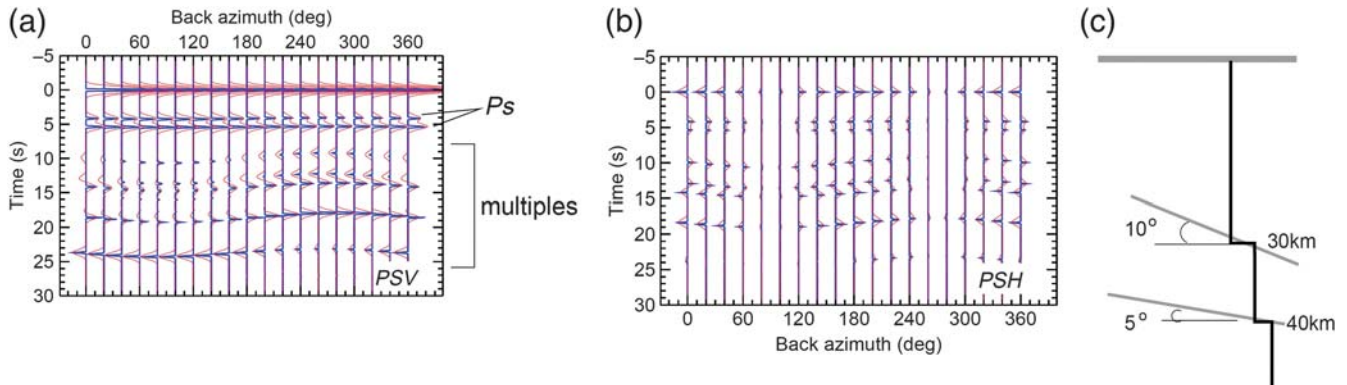
The Kutch rift basin is characterized by a 1–2 km thick top alluvium layer underlain by a 3–4 km thick Jurassic sedimentary layer (Gupta *et al.*, 2001). To examine the ability of

our BPRF technique in delineating thin sedimentary layers, we apply our technique on the broadband waveform data of teleseismic events of magnitude greater than  $M_w$  5, within teleseismic distance ranges of 30°–95° from the Vajepar





**Figure 7.** Synthetic RF and the inversion result for a 5 km thin layer at the surface. (a) Traditional RF for a case of 5 km thick thin layer using time-domain deconvolution (black) and the RF calculated from the best-fitted velocity model (blue) in panel (c). (b) RF estimated using GPSR (black) and the best-fitted velocity model is shown in panel (c) as red. (c) Velocity models calculated from the time-domain RF (blue) and BPRF (red) along with true model (dotted black) and starting model (gray).

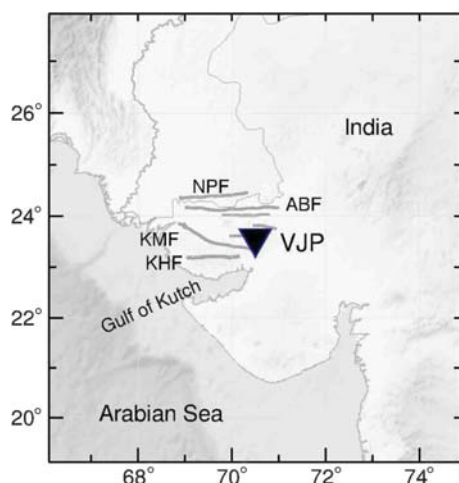


**Figure 8.** Comparison of the synthetic traditional (red) and BP (blue) RFs for dipping layer model along the back azimuth. The (a) radial and (b) transverse components are generated by deconvolving the respective component by the corresponding vertical. The model is shown in (c), where the Moho and the thin layer dip at an angle of  $5^\circ$  and  $10^\circ$ , respectively.

(VJP) station of our Kutch seismic network (Fig. 9). We generated RFs using our modified GPSR algorithm with wedge and nonwedge dictionary and also with time-domain deconvolution techniques (Ligorria and Ammon, 1999). First, we applied instrument corrections to the  $Z-N-E$  components of digital broadband waveforms using pole-zeros information of the sensor. Then, the waveforms were rotated into radial and transverse components using the back-azimuth information. We further rotated the  $Z-R-T$  components into  $P-SV-SH$  components using an incidence angle derived from the minimum  $SV$  energy on the  $P$  component at zero time to have an optimal isolation of  $P-SV-SH$  (Kumar and Kawakatsu, 2011). Once the traces were rotated into  $P$  and  $SV$  components,

we employed source normalization using a time-domain deconvolution technique (Berkhout, 1977). The distance effect due to different source distributions in RFs is corrected by applying moveout to a reference slowness of 6.4 s/deg (Yuan et al., 2006) using the global IASP91 Earth model (Kennett and Engdahl, 1991).

We also treat these rotated waveforms with the basis pursuit deconvolution. The computed radial RFs from our modified GPSR method show clear multiples along with several small amplitude  $PS$  conversions perhaps from other thin layers with large acoustic impedances (Fig. 10). However, the RFs, which are estimated using the iterative time-domain deconvolution method, show broad direct  $P$  and  $PS$  conversions



**Figure 9.** A plot showing one three-component broadband seismograph station. Stations (black solid triangles): VJP, Vajepar. Major faults (lines): ABF, Allah Bund fault; KMF, Kachchh main-land fault; KHF, Katrol Hill fault; NPF, Nagar Parkar fault.

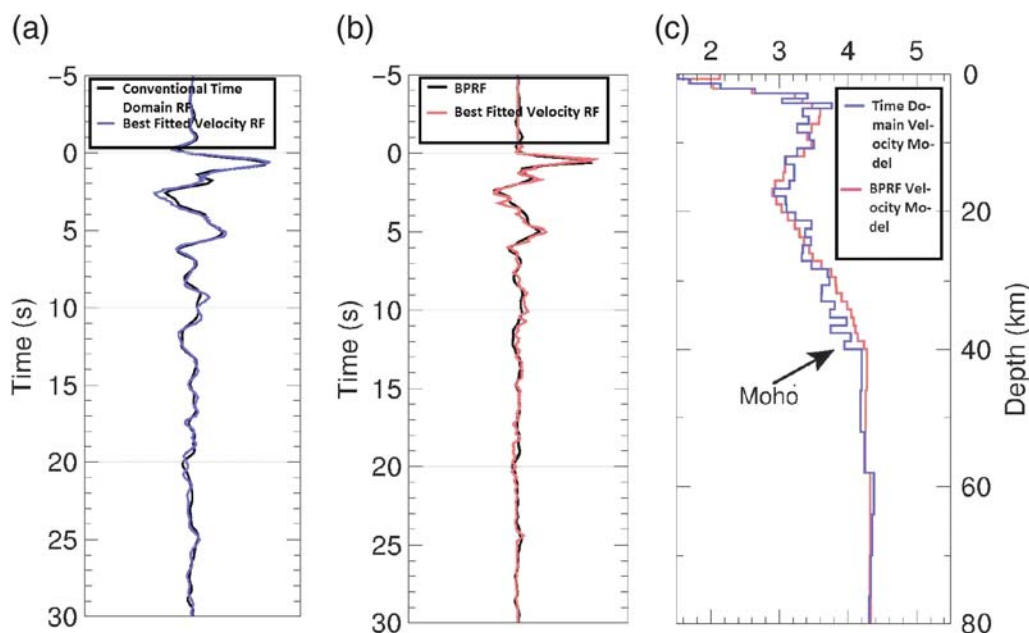
from the Moho, but they do not show any clear splitting of direct  $P$  into small amplitude  $PSS$  conversion (Fig. 10). Even though the conversions are not as clear as they are in the BPRF, they can perhaps be identified there due to subtle changes in the waveforms. The conversion from sediment-basement interface ( $PSS$ ) arrives just after the arrival of direct  $P$  on radial RFs (Fig. 10).

We further demonstrate that the BPRF for observed data can be inverted to obtain a 1D velocity model like traditional RF computed by another deconvolution technique. The

velocity models derived by using the two RF traces are shown in Figure 10c. The derived models are similar in that they resolve a near-surface layer and the Moho appears as a linear gradient rather than a sharp discontinuity. However, the BPRF derived model is smoother than that derived by the traditional RF, which is noisy.

## Discussion and Conclusion

RFs are now fairly popular in imaging crustal and upper-mantle discontinuities. Several methods exist for estimating RFs by time or frequency domain deconvolution. Here, we propose yet another method that makes use of a dipole dictionary and the L1 norm constraint in the deconvolution. Unlike the standard L1 norm constraint on reflectivity that imposes sparseness in reflectivity, the BPD imposes sparseness in layers. Our numerical examples with synthetics as well as real data demonstrate that BPRFs are stable and are able to decipher thin layers in the RFs. Even if thin layer estimation is of no interest in RF studies, the BPRF can be used routinely for obtaining stable RFs. The resolution criterion of BP in case of seismic reflection is discussed by Chung and Lawton (1995) and more recently by Zhang and Castagna (2011) as the tuning thickness of thin beds. Similarly, we define the tuning thickness ( $t_r$ ) of a thin bed resolvable using the  $P$ -to- $S$  converted wave to be  $t_r = \frac{\sqrt{6}}{\pi f_0}$ , in which  $f_0$  is the dominant frequency. One requirement of BPRF is the choice of the regularization weight that determines the relative importance of data and model norms. We demonstrate that a



**Figure 10.** Plot of observed RF stack with inversion for the station VJP in Kutch region. The stack has been generated using 118 individual corresponding RFs. The black wiggles in (a) and (b) are the stack traces after the traditional time-domain deconvolution and GPSR wedge dictionary, respectively. The blue and red wiggles are the best inverted model obtained by the model shown in the same color in panel (c).

theoretical regularization weight that is iteration adaptive is very effective for this purpose.

## Data and Resources

Seismograms used in this study were recorded on one three-component seismograph station of a local seismic network of the Council of Scientific and Industrial Research National Geophysical Research Institute (CSIR-NGRI), Hyderabad, in the Kutch rift zone, Gujarat, India. Data is confidential and is not available to the general public. The BPRF code will be made available to the community.

## Acknowledgments

The work reported in this paper has been supported by the Council of Scientific and Industrial Research twelfth five year plan project at the Council of Scientific and Industrial Research National Geophysical Research Institute (CSIR-NGRI), Hyderabad. M. K. S. was partially supported by the Jackson Chair funds at the Jackson School of Geosciences, University of Texas, Austin.

## References

- Alliney, S., and S. A. Ruzinsky (1994). An algorithm for the minimization of mixed l1 and l2 norms with application to Bayesian estimation, *IEEE Trans. Signal Process.* **42**, 618–627.
- Ammon, C. J. (1991). The isolation of receiver effects from teleseismic *P* waveforms, *Bull. Seismol. Soc. Am.* **81**, 2504–2510.
- Berkhout, A. J. (1977). Least square inverse filtering and wavelet decomposition, *Geophysics* **42**, 1369–1383.
- Biswas, R., and M. K. Sen (2014). Regularization weight estimation in basis pursuit inversion using a Bayesian framework, *SEG Technical Program Expanded Abstracts 2014*, 3237–3241, doi: [10.1190/segam2014-0682.1](https://doi.org/10.1190/segam2014-0682.1).
- Bostock, M. G. (1998). Mantle stratigraphy and evolution of the Slave province, *J. Geophys. Res.* **103**, 21,183–21,200.
- Burdick, L. J., and C. A. Langston (1977). Modeling crustal structure through the use of converted phases in teleseismic body-wave forms, *Bull. Seismol. Soc. Am.* **67**, 677–691.
- Chen, C., D. E. Miller, H. A. Djikpesse, J. B. Haldorsen, and S. Rondenay (2010). Array-conditioned deconvolution of multiple-component teleseismic recordings, *Geophys. J. Int.* **182**, 967–976.
- Chen, S., D. Donoho, and M. Saunders (1998). Atomic decomposition by Basis Pursuit, *SIAM J. Sci. Comput.* **20**, 33–61.
- Chung, H., and D. C. Lawton (1995). Frequency characteristics of seismic reflections from thin beds, *Can. J. Explor. Geophys.* **31**, 32–37.
- Figueiredo, M. A., R. D. Nowak, and S. J. Wright (2007). Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems, *IEEE J. Sel. Top. Signal Process.* **1**, 586–597.
- Frederiksen, A. W., and M. G. Bostock (2000). Modelling teleseismic waves in dipping anisotropic structures, *Geophys. J. Int.* **141**, 401–412.
- Fuchs, K., and G. Müller (1971). Computation of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. Roy. Astron. Soc.* **23**, 417–433.
- Gupta, H. K., T. Harinarayana, M. Kousalya, D. C. Mishra, I. Mohan, N. Purnachandra Rao, P. S. Raju, B. K. Rastogi, P. R. Reddy, and D. Sarkar (2001). Bhuj earthquake of 26 January 2001, *J. Geol. Soc. India* **57**, 275–278.
- Gurrola, H., G. E. Baker, and J. B. Minster (1995). Simultaneous time-domain deconvolution with application to the computation of receiver functions, *Geophys. J. Int.* **120**, 537–543.
- Haldorsen, J., D. Miller, and J. Walsh (1994). Multichannel Wiener deconvolution of vertical seismic profiles, *Geophysics* **59**, 1500–1511.
- Haldorsen, J., D. Miller, and J. Walsh (1995). Walkaway VSP using drill noise as a source, *Geophysics* **60**, 978–997.
- Kennett, B. L. N., and E. R. Engdahl (1991). Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.* **105**, 429–465.
- Kikuchi, M., and H. Kanamori (1982). Inversion of complex body waves, *Bull. Seismol. Soc. Am.* **72**, 491–506.
- Kumar, P., and H. Kawakatsu (2011). Imaging the seismic lithosphere-aesthenosphere boundary of the oceanic plate, *Geochem. Geophys. Geosyst.* **12**, Q01006, doi: [10.1029/2010GC003358](https://doi.org/10.1029/2010GC003358).
- Ligorria, J. P., and C. J. Ammon (1999). Iterative deconvolution and receiver-function estimation, *Bull. Seismol. Soc. Am.* **89**, 1395–1400.
- Owens, T. J., G. Zandt, and S. R. Taylor (1984). Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee: A detailed analysis of broadband teleseismic *P* waveforms, *J. Geophys. Res.* **89**, 7783–7795.
- Phinney, R. A. (1964). Structure of the Earth's crust from spectral behavior of long-period body waves, *J. Geophys. Res.* **69**, 2997–3017.
- Puryear, C. I., and J. P. Castagna (2008). Layer-thickness determination and stratigraphic interpretation using spectral inversion: Theory and application, *Geophysics* **73**, no. 2, R37–R48.
- Sen, M. K., and P. L. Stoffa (2013). Global optimization methods in geophysical inversion, Cambridge University Press, Cambridge, United Kingdom.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation, *Geophysics* **49**, 1259–1266.
- Vinnik, L. (1977). Detection of waves converted from *P* to *SV* in the mantle, *Phys. Earth Planet. In.* **15**, 39–45.
- Yuan, X., R. Kind, X. Li, and R. Wang (2006). The *S* receiver functions: Synthetics and data example, *Geophys. J. Int.* **165**, no. 2, 555–564, doi: [10.1111/j.1365-246X.2006.02885.x](https://doi.org/10.1111/j.1365-246X.2006.02885.x).
- Zhang, R., and J. Castagna (2011). Seismic sparse-layer reflectivity inversion using basis pursuit decomposition, *Geophysics* **76**, R147–R158.
- Zhang, R., M. K. Sen, and S. Srinivasan (2013). A prestack basis pursuit seismic inversion, *Geophysics* **78**, R1–R11.

National Geophysical Research Institute (CSIR-NGRI)  
Uppal Road  
Hyderabad  
Andhra Pradesh 500007, India  
msentx@gmail.com  
mrinal@ig.utexas.edu  
reetam.biswas@gmail.com  
prantikmandal62@gmail.com  
prakashk@ngri.res.in

Manuscript received 3 January 2014;

Published Online 4 November 2014